

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: April 4, 2003, 16:35:42 ; Search time 1467 Seconds
(without alignments)
2142.538 Million cell updates/sec

Title: US-09-541-462B-2

Perfect score: 616

Sequence: 1 MAAAMDVDTPSGTNSGAGKK.....KTRQVCPLDNREWEFQKYGH 108

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+_p2n.model -DEV=xlp

-

Q=/cgn2_1/USPTO_spool/US09541462/runat_31032003_090927_7154/app_query.fasta_1.26
3

-DB=GenEmbl -QFMT=fastap -SUFFIX=std.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09541462@CGN_1_1_3745@runat_31032003_090927_7154 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl:*

1: gb_ba:*

2: gb_htg:*

3: gb_in:*

4: gb_om:*

5: gb_ov:*

6: gb_pat:*

7: gb_ph:*

8: gb_pl:*

9: gb_pr:*

10: gb_ro:*

```

11:  gb_sts:*
12:  gb_sy:*
13:  gb_un:*
14:  gb_vi:*
15:  em_ba:*
16:  em_fun:*
17:  em_hum:*
18:  em_in:*
19:  em_mu:*
20:  em_om:*
21:  em_or:*
22:  em_ov:*
23:  em_pat:*
24:  em_ph:*
25:  em_pl:*
26:  em_ro:*
27:  em_sts:*
28:  em_un:*
29:  em_vi:*
30:  em_htg_hum:*
31:  em_htg_inv:*
32:  em_htg_other:*
33:  em_htg_mus:*
34:  em_htg_pln:*
35:  em_htg_rod:*
36:  em_htg_mam:*
37:  em_htg_vrt:*
38:  em_sy:*
39:  em_htgo_hum:*
40:  em_htgo_mus:*
41:  em_htgo_other:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result | Query | | | | | Description |
|--------|-------|-------|-------|--------|------------|--------------------|
| | No. | Score | Match | Length | DB | |
| ID | | | | | | |
| 1 | 616 | 100.0 | 327 | 9 | AF142059 | AF142059 Homo sapi |
| 2 | 616 | 100.0 | 504 | 10 | AF140599 | AF140599 Mus muscu |
| 3 | 616 | 100.0 | 508 | 9 | AF140598 | AF140598 Homo sapi |
| 4 | 616 | 100.0 | 544 | 9 | BC001466 | BC001466 Homo sapi |
| 5 | 616 | 100.0 | 554 | 9 | BC017370 | BC017370 Homo sapi |
| 6 | 616 | 100.0 | 1616 | 10 | BC027396 | BC027396 Mus muscu |
| 7 | 604 | 98.1 | 1933 | 5 | AY027936 | AY027936 Salmo sal |
| 8 | 584 | 94.8 | 497 | 9 | HUMYQ60A05 | AF085906 Homo sapi |
| c 9 | 573 | 93.0 | 5347 | 6 | AX281690 | AX281690 Sequence |
| c 10 | 554 | 89.9 | 3484 | 9 | HTEST | X73608 H.sapiens m |
| 11 | 545 | 88.5 | 306 | 9 | AY099360 | AY099360 Homo sapi |
| 12 | 535 | 86.9 | 1181 | 3 | AY119265 | AY119265 Drosophil |
| 13 | 515.5 | 83.7 | 635 | 8 | AY086913 | AY086913 Arabidops |
| 14 | 515.5 | 83.7 | 666 | 8 | AY072430 | AY072430 Arabidops |
| 15 | 514.5 | 83.5 | 357 | 8 | AY052401 | AY052401 Arabidops |
| 16 | 514.5 | 83.5 | 544 | 8 | AY114719 | AY114719 Arabidops |

| | | | | | | |
|------|-------|------|--------|----|-----------|--------------------|
| 17 | 501.5 | 81.4 | 3208 | 6 | AX212267 | AX212267 Sequence |
| 18 | 492.5 | 80.0 | 361 | 3 | AB077287 | AB077287 Caenorhab |
| 19 | 469 | 76.1 | 324 | 8 | AF179228 | AF179228 Schizosac |
| 20 | 464.5 | 75.4 | 46630 | 3 | DMC115C2 | AL031581 Drosophil |
| c 21 | 464.5 | 75.4 | 102227 | 2 | AC020129 | AC020129 Drosophil |
| 22 | 464.5 | 75.4 | 161668 | 3 | AC104602 | AC104602 Drosophil |
| 23 | 464.5 | 75.4 | 299970 | 3 | AE003418 | AE003418 Drosophil |
| c 24 | 433 | 70.3 | 433 | 11 | G27926 | G27926 human STS S |
| 25 | 428 | 69.5 | 40197 | 3 | CBRG33P21 | AC084553 Caenorhab |
| 26 | 425.5 | 69.1 | 171237 | 2 | AC123343 | AC123343 Rattus no |
| 27 | 411.5 | 66.8 | 40387 | 8 | SPAC23H4 | Z98977 S.pombe chr |
| 28 | 410.5 | 66.6 | 104787 | 9 | AC109638 | AC109638 Homo sapi |
| c 29 | 402.5 | 65.3 | 39874 | 3 | CEZK287 | Z70757 Caenorhabdi |
| 30 | 400.5 | 65.0 | 553 | 3 | AY070810 | AY070810 Drosophil |
| 31 | 400.5 | 65.0 | 750 | 3 | AF218290 | AF218290 Drosophil |
| c 32 | 400.5 | 65.0 | 3185 | 3 | AY061302 | AY061302 Drosophil |
| 33 | 400.5 | 65.0 | 95294 | 2 | AC019742 | AC019742 Drosophil |
| 34 | 400.5 | 65.0 | 167977 | 3 | AC010010 | AC010010 Drosophil |
| 35 | 400.5 | 65.0 | 310120 | 3 | AE003468 | AE003468 Drosophil |
| 36 | 385 | 62.5 | 1255 | 8 | SCYOL134C | Z74876 S.cerevisia |
| 37 | 385 | 62.5 | 12805 | 8 | SC128XV | X95465 S.cerevisia |
| c 38 | 378 | 61.4 | 164399 | 3 | PFMAL3P6 | Z98551 Plasmodium |
| 39 | 352.5 | 57.2 | 173346 | 9 | AC112184 | AC112184 Homo sapi |
| 40 | 352.5 | 57.2 | 175561 | 9 | AC012636 | AC012636 Homo sapi |
| c 41 | 335.5 | 54.5 | 79677 | 8 | ATT21C14 | AL138639 Arabidops |
| c 42 | 327.5 | 53.2 | 177444 | 2 | AC024617 | AC024617 Homo sapi |
| 43 | 314 | 51.0 | 209982 | 3 | CNS06C8G | AL391737 chromosom |
| 44 | 292 | 47.4 | 342 | 9 | AF142060 | AF142060 Homo sapi |
| c 45 | 292 | 47.4 | 418 | 6 | AX332170 | AX332170 Sequence |

ALIGNMENTS

RESULT 1
 AF142059

LOCUS AF142059 327 bp mRNA linear PRI 14-JUL-1999
 DEFINITION Homo sapiens RING finger protein (ROC1) mRNA, complete cds.
 ACCESSION AF142059
 VERSION AF142059.1 GI:4809215
 KEYWORDS .
 SOURCE Homo sapiens.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 327)
 AUTHORS Ohta,T., Michel,J.J., Schottelius,A.J. and Xiong,Y.
 TITLE ROC1, a homolog of APC11, represents a family of cullin partners
 with an associated ubiquitin ligase activity
 JOURNAL Mol. Cell 3 (4), 535-541 (1999)
 MEDLINE 99247022
 PUBMED 10230407
 REFERENCE 2 (bases 1 to 327)
 AUTHORS Ohta,T., Michel,J. and Xiong,Y.
 TITLE Direct Submission
 JOURNAL Submitted (10-APR-1999) Lineberger Comprehensive Cancer Center,
 University of North Carolina at Chapel Hill, Mason Farm Rd. and
 Manning Dr., Chapel Hill, NC 27599-7295, USA

```

FEATURES          Location/Qualifiers
source           1. .327
                  /organism="Homo sapiens"
                  /db_xref="taxon:9606"
gene            1. .327
                  /gene="ROC1"
CDS             1. .327
                  /gene="ROC1"
                  /note="Cullin partnering protein with associated ubiquitin
                  ligase activity"
                  /codon_start=1
                  /product="RING finger protein"
                  /protein_id="AAD30146.1"
                  /db_xref="GI:4809216"
                  /translation="MAAAMDVDTPSGTNSGAGKKRFEVKWNAVALWA
                  DIVVDNCAI
                  CRNHIMDLCIECQANQASATSEECTVAWGVCNHFHCISRWLKTRQVCPLDNREWE
                  FQKYGH"
BASE COUNT      85 a      75 c      94 g      73 t
ORIGIN

```

Alignment Scores:

| | | | |
|------------------------|----------|---------------|-----|
| Pred. No.: | 5.64e-59 | Length: | 327 |
| Score: | 616.00 | Matches: | 108 |
| Percent Similarity: | 100.00% | Conservative: | 0 |
| Best Local Similarity: | 100.00% | Mismatches: | 0 |
| Query Match: | 100.00% | Indels: | 0 |
| DB: | 9 | Gaps: | 0 |

US-09-541-462B-2 (1-108) x AF142059 (1-327)

| | | | |
|----|-----|--|-----|
| Qy | 1 | MetAlaAlaAlaMetAspValAspThrProSerGlyThrAsnSerGlyAlaGlyLysLys | 20 |
| | | | |
| Db | 1 | ATGGCGGCAGCGATGGATGTGGATACCCCGAGCGGCACCAACAGCGCGCGGGCAAGAAG | 60 |
| Qy | 21 | ArgPheGluValLysLysTrpAsnAlaValAlaLeuTrpAlaTrpAspIleValValAsp | 40 |
| | | | |
| Db | 61 | CGCTTGAAAGTAAAAAGTGGATGCAGTAGCCCTCTGGGCCTGGATATTGTGGTTGAT | 120 |
| Qy | 41 | AsnCysAlaIleCysArgAsnHisIleMetAspLeuCysIleGluCysGlnAlaAsnGln | 60 |
| | | | |
| Db | 121 | AACTGTGCCATCTGCAGGAACCACATTATGGATCTTGATAGAAATGTCAGCTAACAG | 180 |
| Qy | 61 | AlaSerAlaThrSerGluGluCysThrValAlaTrpGlyValCysAsnHisAlaPheHis | 80 |
| | | | |
| Db | 181 | GCGTCCGCTACTCAGAAGAGTGTACTGTCGATGGGAGTCTGTAACCATGCTTTCAC | 240 |
| Qy | 81 | PheHisCysIleSerArgTrpLeuLysThrArgGlnValCysProLeuAspAsnArgGlu | 100 |
| | | | |
| Db | 241 | TTCCACTGCATCTCGCTGGCTAAAACACGACAGGTGTCCATTGGACAACAGAGAG | 300 |
| Qy | 101 | TrpGluPheGlnLysTyrGlyHis | 108 |
| | | | |
| Db | 301 | TGGGAATTCCAAAAGTATGGGCAC | 324 |

RESULT 2

AF140599

LOCUS AF140599 504 bp mRNA linear ROD 11-MAY-1999

DEFINITION Mus musculus ring-box protein 1 (Rbx1) mRNA, complete cds.
 ACCESSION AF140599
 VERSION AF140599.1 GI:4769005
 KEYWORDS
 SOURCE Mus musculus.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 504)
 REFERENCE
 AUTHORS Kamura,T., Koepp,D.M., Conrad,M.N., Skowyra,D., Moreland,R.J.,
 Iliopoulos,O., Lane,W.S., Kaelin,W.G. Jr., Elledge,S.J.,
 Conaway,R.C., Harper,J.W. and Conaway,J.W.
 TITLE Rbx1, a component of the VHL tumor suppressor complex and SCF
 ubiquitin ligase
 JOURNAL Science 284 (5414), 657-661 (1999)
 MEDLINE 99234320
 PUBMED 10213691
 REFERENCE 2 (bases 1 to 504)
 AUTHORS Kamura,T., Lane,W.S., Conaway,R.C. and Conaway,J.W.
 TITLE Direct Submission
 JOURNAL Submitted (05-APR-1999) Prog. Molec. Cell Biol., HHMI, Oklahoma
 Med. Res. Fdn., 825 NE 13th St, Oklahoma City, OK 73104, USA
 FEATURES Location/Qualifiers
 source 1. .504
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 gene 1. .504
 /gene="Rbx1"
 CDS 18. .344
 /gene="Rbx1"
 /note="component of VHL tumor suppressor complex and SCF
 ubiquitin ligase"
 /codon_start=1
 /product="ring-box protein 1"
 /protein_id="AAD29716.1"
 /db_xref="GI:4769006"
 /translation="MAAAMDVDTPSGTNSGAGKKRFEVKKWNAVALWAWDIVVDNCAI
 CRNHIMDLCIECQANQASATSEECTVAWGVCNHAFFHCISRWLKTRQVCPLDNREWE
 FQKYGH"
 BASE COUNT 117 a 107 c 137 g 143 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 8.92e-59 Length: 504
 Score: 616.00 Matches: 108
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 10 Gaps: 0

US-09-541-462B-2 (1-108) x AF140599 (1-504)

Qy 1 MetAlaAlaAlaMetAspValAspThrProSerGlyThrAsnSerGlyAlaGlyLysLys 20
 ||||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 18 ATGGCGGCGGCGATGGATGTGGATACCCCCAGCGGCACCAACAGCGGCGGGCAAGAAG 77
 Qy 21 ArgPheGluValLysLysTrpAsnAlaValAlaLeuTrpAlaTrpAspIleValValAsp 40

```

Db      78 CGCTTGAAGTTAAAAAGTGGAAATGCAGTGGCCCTCTGGCCTGGGACATTGTGGTGAT 137
Qy      41 AsnCysAlaIleCysArgAsnHisIleMetAspLeuCysIleGluCysGlnAlaAsnGln 60
Db      138 AACTGTGCCATCTGCAGGAACCACATTATGGATCTTGTATCGAATGTCAGGCCAACCA 197
Qy      61 AlaSerAlaThrSerGluGluCysThrValAlaTrpGlyValCysAsnHisAlaPheHis 80
Db      198 GCGTCAGCTACTTCCGAAGAGTGTACGGTTGCATGGGGAGTCTGCAACCATGCTTTCAT 257
Qy      81 PheHisCysIleSerArgTrpLeuLysThrArgGlnValCysProLeuAspAsnArgGlu 100
Db      258 TTCCACTGCATCTCTCGATGGCTAAAACGAGGCAGGTGTCCGTTGGACAACAGAGAG 317
Qy      101 TrpGluPheGlnLysTyrGlyHis 108
Db      318 TGGGAGTTCCAGAAGTATGGGCAT 341

```

RESULT 3

AF140598

LOCUS AF140598 508 bp mRNA linear PRI 11-MAY-1999
 DEFINITION Homo sapiens ring-box protein 1 (RBX1) mRNA, complete cds.
 ACCESSION AF140598
 VERSION AF140598.1 GI:4769003
 KEYWORDS .
 SOURCE Homo sapiens.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 508)
 AUTHORS Kamura,T., Koepp,D.M., Conrad,M.N., Skowyra,D., Moreland,R.J.,
 Iliopoulos,O., Lane,W.S., Kaelin,W.G. Jr., Elledge,S.J.,
 Conaway,R.C., Harper,J.W. and Conaway,J.W.
 TITLE Rbx1, a component of the VHL tumor suppressor complex and SCF
 ubiquitin ligase
 JOURNAL Science 284 (5414), 657-661 (1999)
 MEDLINE 99234320
 PUBMED 10213691
 REFERENCE 2 (bases 1 to 508)
 AUTHORS Kamura,T., Lane,W.S., Conaway,R.C. and Conaway,J.W.
 TITLE Direct Submission
 JOURNAL Submitted (05-APR-1999) Prog. Molec. Cell Biol., HHMI, Oklahoma
 Med. Res. Fdn., 825 NE 13th St, Oklahoma City, OK 73104, USA
 FEATURES Location/Qualifiers
 source 1. .508
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 gene 1. .508
 /gene="RBX1"
 CDS 7. .333
 /gene="RBX1"
 /note="ring finger-like protein; component of VHL tumor
 suppressor complex and SCF ubiquitin ligase"
 /codon_start=1
 /product="ring-box protein 1"
 /protein_id="AAD29715.1"

/db_xref="GI:4769004"
/translation="MAAAMDVDPSTGAGKKRFEVKKWNAVALWAWDIVVDNCAI
CRNHIMDLCIECQANQASATSEECTVAWGVCNHAFHFCISRWLKTRQVCPLDNREWE
FQKYGH"

BASE COUNT 126 a 106 c 124 g 152 t
ORIGIN

Alignment Scores:

| | | | |
|------------------------|----------|---------------|-----|
| Pred. No.: | 8.99e-59 | Length: | 508 |
| Score: | 616.00 | Matches: | 108 |
| Percent Similarity: | 100.00% | Conservative: | 0 |
| Best Local Similarity: | 100.00% | Mismatches: | 0 |
| Query Match: | 100.00% | Indels: | 0 |
| DB: | 9 | Gaps: | 0 |

US-09-541-462B-2 (1-108) x AF140598 (1-508)

| | | |
|----|--|--|
| Qy | 1 Met Ala Ala Ala Met Asp Val Asp Thr Pro Ser Gly Thr Asn Ser Gly Ala Gly Lys Lys 20 | |
| Db | 7 ATGGCGGCAGCGATGGATGTGGATACCCCGAGCGGCACCAACAGCGCGCGGGCAAGAAG 66 | |
| Qy | 21 Arg Phe Glu Val Lys Lys Trp Asn Ala Val Ala Leu Trp Ala Trp Asp Ile Val Val Asp 40 | |
| Db | 67 CGCTTTGAAGTGAAGAAAGTGGAAATGCAGTAGCCCTCTGGCCTGGGATATTGTGGTTGAT 126 | |
| Qy | 41 Asn Cys Ala Ile Cys Arg Asn His Ile Met Asp Leu Cys Ile Glu Cys Gln Ala Asn Gln 60 | |
| Db | 127 AACTGTGCCATCTGCAGGAACCACATTATGGATCTTGCATAGAATGTCAAGCTAACAG 186 | |
| Qy | 61 Ala Ser Ala Thr Ser Glu Glu Cys Thr Val Ala Trp Gly Val Cys Asn His Ala Phe His 80 | |
| Db | 187 GCGTCCGCTACTTCAGAAGAGTGTACTGTCGATGGGAGTCTGTAACCATGCTTTCAC 246 | |
| Qy | 81 Phe His Cys Ile Ser Arg Trp Leu Lys Thr Arg Gln Val Cys Pro Leu Asp Asn Arg Glu 100 | |
| Db | 247 TTCCACTGCATCTCTCGCTGGCTAAAACACGACAGGTGTCCATTGGACAAACAGAGAG 306 | |
| Qy | 101 Trp Glu Phe Gln Lys Tyr Gly His 108 | |
| Db | 307 TGGGAATTCCAAAAGTATGGGCAC 330 | |

RESULT 4

BC001466

LOCUS BC001466 544 bp mRNA linear PRI 12-JUL-2001
DEFINITION Homo sapiens, ring-box 1, clone MGC:1481 IMAGE:3138751, mRNA, complete cds.
ACCESSION BC001466
VERSION BC001466.1 GI:12655214
KEYWORDS MGC.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 544)
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (12-DEC-2000) National Institutes of Health, Mammalian

Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>
COMMENT Contact: MGC help desk
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC), Gaithersburg, Maryland;
Web site: <http://www.nisc.nih.gov/>
Contact: nisc_mgc@nhgri.nih.gov
Shevchenko, Y., Wetherby, K.D., Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W., Bouffard, G.G., Brinkley, C., Brooks, S., Dietrich, N.L., Guan, X., Gupta, J., Ho, S.-L., Karlins, E., Legaspi, R., Lim, M., Maduro, Q.L., Masiello, C., Mastrian, S.D., McCloskey, J.C., McDowell, J., Pearson, R., Snyder, B., Stantripop, S., Thomas, P.J., Tiongson, E.E., Touchman, J.W., Tsurgeon, C., Vogt, J.L., Walker, M.A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAL Plate: 4 Row: j Column: 15.

FEATURES Location/Qualifiers
source 1. .544
/organism="Homo sapiens"
/db_xref="LocusID:9978"
/db_xref="taxon:9606"
/clone="MGC:1481 IMAGE:3138751"
/tissue_type="Placenta, choriocarcinoma"
/clone_lib="NIH_MGC_21"
/lab_host="DH10B-R"
/note="Vector: pOTB7"
CDS 15. .341
/codon_start=1
/product="ring-box 1"
/protein_id="AAH01466.1"
/db_xref="GI:12655215"
/translation="MAAADVDTPSGTNSGAGKKRFEVKWNAVALWAWDIVVDNCAI
CRNHIMDLCIECQANQASATSEECTVAWGVCNHFHCISRWLKTRQVCPLDNREWE
FQKYGH"

BASE COUNT 157 a 107 c 129 g 151 t
ORIGIN

Alignment Scores:

| | | | |
|------------------------|----------|---------------|-----|
| Pred. No.: | 9.67e-59 | Length: | 544 |
| Score: | 616.00 | Matches: | 108 |
| Percent Similarity: | 100.00% | Conservative: | 0 |
| Best Local Similarity: | 100.00% | Mismatches: | 0 |
| Query Match: | 100.00% | Indels: | 0 |
| DB: | 9 | Gaps: | 0 |

US-09-541-462B-2 (1-108) x BC001466 (1-544)

Qy 1 MetAlaAlaAlaMetAspValAspThrProSerGlyThrAsnSerGlyAlaGlyLysLys 20

```

Db      |||||||ATGGCGGCAGCGATGGATGTGGATACCCCGAGCGGCACCAACAGCGCGCGGGCAAGAAG 74
Qy      21 ArgPheGluValLysLysTrpAsnAlaValAlaLeuTrpAlaTrpAspIleValValAsp 40
Db      75 CGCTTTGAAGTGAAAAAGTGGAAATGCAGTAGCCCTCTGGGCCTGGGATATTGTGGTTGAT 134
Qy      41 AsnCysAlaIleCysArgAsnHisIleMetAspLeuCysIleGluCysGlnAlaAsnGln 60
Db      135 AACTGTGCCATCTGCAGGAACCACATTATGGATCTTGCATAGAATGTCAAGCTAACAG 194
Qy      61 AlaSerAlaThrSerGluGluCysThrValAlaTrpGlyValCysAsnHisAlaPheHis 80
Db      195 GCGTCCGCTACTTCAGAAAGAGTGTACTGTCGATGGGAGTCTGTAACCATGCTTTCAC 254
Qy      81 PheHisCysIleSerArgTrpLeuLysThrArgGlnValCysProLeuAspAsnArgGlu 100
Db      255 TTCCACTGCATCTCGCTGGCTAAAACACGACAGGTGTCCATTGGACAACAGAGAG 314
Qy      101 TrpGluPheGlnLysTyrGlyHis 108
Db      315 TGGGAATTCCAAAAGTATGGGCAC 338

```

RESULT 5

BC017370

LOCUS BC017370 554 bp mRNA linear PRI 14-NOV-2001
DEFINITION Homo sapiens, ring-box 1, clone MGC:13357 IMAGE:4065797, mRNA, complete cds.
ACCESSION BC017370
VERSION BC017370.1 GI:16924201
KEYWORDS MGC.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 554)
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (13-NOV-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>
COMMENT Contact: MGC help desk
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: CLONTECH Laboratories, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Institute for Systems Biology
<http://www.systemsbiology.org>
contact: amadan@systemsbiology.org
Anup Madan, Jessica Fahey, Erin Helton, Mark Ketteman, Anuradha
Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAL Plate: 19 Row: c Column: 17

This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 7657507.

FEATURES Location/Qualifiers
source 1. .554
/organism="Homo sapiens"
/db_xref="LocusID:9978"
/db_xref="taxon:9606"
/clone="MGC:13357 IMAGE:4065797"
/tissue_type="Brain, glioblastoma"
/clone_lib="NIH_MGC_57"
/lab_host="DH10B"
/note="Vector: pDNR-LIB"
CDS 23. .349
/codon_start=1
/product="ring-box 1"
/protein_id="AAH17370.1"
/db_xref="GI:16924202"
/translation="MAAAMDVDTPSGTNSGAGKKRFEVKKWNAVALWAUDIVVDNCAI
CRNHIMDLCIECQANQASATSEECTVAVGVCNHAFHFCISRWLKTRQVCPLDNREWE
FQKYGH"
BASE COUNT 157 a 107 c 133 g 157 t
ORIGIN

Alignment Scores:

| | | | |
|------------------------|----------|---------------|-----|
| Pred. No.: | 9.86e-59 | Length: | 554 |
| Score: | 616.00 | Matches: | 108 |
| Percent Similarity: | 100.00% | Conservative: | 0 |
| Best Local Similarity: | 100.00% | Mismatches: | 0 |
| Query Match: | 100.00% | Indels: | 0 |
| DB: | 9 | Gaps: | 0 |

US-09-541-462B-2 (1-108) x BC017370 (1-554)

| | | | |
|----|-----|---|-----|
| Qy | 1 | Met Ala Ala Ala Met Asp Val Asp Thr Pro Ser Gly Thr Asn Ser Gly Ala Gly Lys Lys | 20 |
| Db | 23 | ATGGCGGCAGCGATGGATGTGGATACCCCGAGCGGCACCAACAGCGCGCGGGCAAGAAG | 82 |
| Qy | 21 | Arg Phe Glu Val Lys Lys Trp Asn Ala Val Ala Leu Trp Ala Trp Asp Ile Val Val Asp | 40 |
| Db | 83 | CGCTTGAAAGTGGAAATGCAGTAGCCCTCTGGGCCTGGGATATTGTGGTTGAT | 142 |
| Qy | 41 | Asn Cys Ala Ile Cys Arg Asn His Ile Met Asp Leu Cys Ile Glu Cys Gln Ala Asn Gln | 60 |
| Db | 143 | AACTGTGCCATCTGCAGGAACCACATTATGGATCTTGCATAGAATGTCAAGCTAACCAAG | 202 |
| Qy | 61 | Ala Ser Ala Thr Ser Glu Glu Cys Thr Val Ala Trp Gly Val Cys Asn His Ala Phe His | 80 |
| Db | 203 | GCGTCCGCTACTTCAGAAGAGTGTACTGTCGCATGGGAGTCTGTAACCAGCTTTAC | 262 |
| Qy | 81 | Phe His Cys Ile Ser Arg Trp Leu Lys Thr Arg Gln Val Cys Pro Leu Asp Asn Arg Glu | 100 |
| Db | 263 | TTCCACTGCATCTCGCTGGCTAAAACACGACAGGTGTCCATTGGACAAACAGAGAG | 322 |
| Qy | 101 | Trp Glu Phe Gln Lys Tyr Gly His | 108 |
| Db | 323 | TGGGAATTCCAAAAGTATGGGCAC | 346 |

RESULT 6

BC027396

LOCUS BC027396 1616 bp mRNA linear ROD 07-AUG-2002
DEFINITION Mus musculus, ring-box 1, clone MGC:35907 IMAGE:4952242, mRNA, complete cds.
ACCESSION BC027396
VERSION BC027396.1 GI:20072075
KEYWORDS MGC.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 1616)
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (04-APR-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>
COMMENT Contact: MGC help desk
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Jeffrey Green M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center
Center code: BCM-HGSC
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
Contact: amg@bcm.tmc.edu
Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Hale, S.M., Yoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M., Richards, S., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Series: IRAK Plate: 60 Row: 1 Column: 8

This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 9790190.

FEATURES Location/Qualifiers
source 1. .1616
/organism="Mus musculus"
/db_xref="taxon:10090"
/map="FVB/N"
/clone="MGC:35907 IMAGE:4952242"
/tissue_type="Mammary tumor. C3(1)-Tag model. Infiltrating ductal carcinoma. 5 month old virgin mouse."
/clone_lib="NCI_CGAP_Mam6"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
CDS 7. .333
/codon_start=1
/product="ring-box 1"
/protein_id="AAH27396.1"
/db_xref="GI:20072076"
/db_xref="LocusID:9978"
/translation="MAAAMDVDTPSGTNSGAGKKRFEVKKWNAVALWAUDIVVDNCAI

CRNHIMDLCIECQANQASATSEECTVAWGVCNHAFHFHCISRWLKTRQVCPLDNREWE
FQKYGH"

BASE COUNT 453 a 330 c 407 g 426 t
ORIGIN

Alignment Scores:

| | | | |
|------------------------|----------|---------------|------|
| Pred. No.: | 3.06e-58 | Length: | 1616 |
| Score: | 616.00 | Matches: | 108 |
| Percent Similarity: | 100.00% | Conservative: | 0 |
| Best Local Similarity: | 100.00% | Mismatches: | 0 |
| Query Match: | 100.00% | Indels: | 0 |
| DB: | 10 | Gaps: | 0 |

US-09-541-462B-2 (1-108) x BC027396 (1-1616)

| | |
|----|--|
| Qy | 1 Met Ala Ala Ala Met Asp Val Asp Thr Pro Ser Gly Thr Asn Ser Gly Ala Gly Lys Lys 20 |
| Db | 7 ATGGCGGCCGCGATGGATGTGGATACCCCCAGCGGCACCAACAGCGCGCGGGCAAGAAG 66 |
| Qy | 21 Arg Phe Glu Val Lys Lys Trp Asn Ala Val Ala Leu Trp Ala Trp Asp Ile Val Val Asp 40 |
| Db | 67 CGCTTGAAAGTTAAAAAGTGGATGCAGTGGCCCTCTGGCCTGGGACATTGTGGTTGAT 126 |
| Qy | 41 Asn Cys Ala Ile Cys Arg Asn His Ile Met Asp Leu Cys Ile Glu Cys Gln Ala Asn Gln 60 |
| Db | 127 AACTGTGCCATCTGCAGGAACCACATTATGGATCTTGTATCGAATGTCAGGCCAACAG 186 |
| Qy | 61 Ala Ser Ala Thr Ser Glu Glu Cys Thr Val Ala Trp Gly Val Cys Asn His Ala Phe His 80 |
| Db | 187 GCGTCAGCTACTTCCGAAGAGTGTACGGTTGCATGGGAGTCTGCAACCATGCTTTCAT 246 |
| Qy | 81 Phe His Cys Ile Ser Arg Trp Leu Lys Thr Arg Gln Val Cys Pro Leu Asp Asn Arg Glu 100 |
| Db | 247 TTCCACTGCATCTCGATGGCTAAAACGAGGCAGGTGTCCGTTGGACAACAGAGAG 306 |
| Qy | 101 Trp Glu Phe Gln Lys Tyr Gly His 108 |
| Db | 307 TGGGAGTTCCAGAACTATGGGCAT 330 |

RESULT 7

AY027936

| | | | | | |
|------------|--|---------|------|--------|-----------------|
| LOCUS | AY027936 | 1933 bp | mRNA | linear | VRT 19-MAR-2002 |
| DEFINITION | Salmo salar hyperosmotic protein 21 mRNA, complete cds. | | | | |
| ACCESSION | AY027936 | | | | |
| VERSION | AY027936.1 GI:19067883 | | | | |
| KEYWORDS | . | | | | |
| SOURCE | Salmo salar. | | | | |
| ORGANISM | Salmo salar Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Protacanthopterygii; Salmoniformes; Salmonidae; Salmo. | | | | |
| REFERENCE | 1 (bases 1 to 1933) | | | | |
| AUTHORS | Pan, F., Zarate, J. and Bradley, T.M. | | | | |
| TITLE | A homolog of the E3 ubiquitin ligase Rbx1 is induced during hyperosmotic stress of salmon | | | | |
| JOURNAL | Am. J. Physiol. Regul. Integr. Comp. Physiol. 282 (2002) In press | | | | |
| REFERENCE | 2 (bases 1 to 1933) | | | | |

AUTHORS Bradley, T.M. and Pan, F.
 TITLE Direct Submission
 JOURNAL Submitted (22-FEB-2001) Fisheries, Animal and Veterinary Science,
 University of Rhode Island, Bldg.14, East Farm, Route 108,
 Kingston, RI 02881, USA
 FEATURES Location/Qualifiers
 source 1. .1933
 /organism="Salmo salar"
 /db_xref="taxon:8030"
 /tissue_type="gill"
 CDS 1135. .1704
 /note="SHOP21; putative E3 ubiquitin ligase; similar to
 human and murine RBX1 protein"
 /codon_start=1
 /product="hyperosmotic protein 21"
 /protein_id="AAK29182.1"
 /db_xref="GI:19067884"
 /translation="MSEGVPSPVCPSTLAQAHQSPHRLPMEGQAAQAAGSTNVCWLR
 LKASADHЛИHVALRILCNVMHSHIVHLVPNSARAIVNMAAMDVDTPSATNSGASKKR
 FEVKKWNALWAWDIVVDNCAICRNHIMDLCIECQANQASATSEECTVAWGVCNHAF
 HFHCISRWLKTRQVCPLDNREWEFQKYGH"
 BASE COUNT 518 a 466 c 443 g 506 t
 ORIGIN

Alignment Scores:

| | | | |
|------------------------|---------|---------------|------|
| Pred. No.: | 7.7e-57 | Length: | 1933 |
| Score: | 604.00 | Matches: | 106 |
| Percent Similarity: | 98.15% | Conservative: | 0 |
| Best Local Similarity: | 98.15% | Mismatches: | 2 |
| Query Match: | 98.05% | Indels: | 0 |
| DB: | 5 | Gaps: | 0 |

US-09-541-462B-2 (1-108) x AY027936 (1-1933)

| | | |
|----|------|---|
| Qy | 1 | Met Ala Ala Ala Met Asp Val Asp Thr Pro Ser Gly Thr Asn Ser Gly Ala Gly Lys Lys 20 |
| Db | 1378 | ATGGCGGCAGCGATGGATGTTGATACCCCAAGCGCCACGAATAGTGGAGCAAGCAAGAAA 1437 |
| Qy | 21 | Arg Phe Glu Val Lys Lys Trp Asn Ala Val Ala Leu Trp Ala Trp Asp Ile Val Val Asp 40 |
| Db | 1438 | CGTTTGAAAGTGAAGAAGTGGAAATGCAGTGGCACTTGGGCCTGGGACATTGTGGTGGAC 1497 |
| Qy | 41 | Asn Cys Ala Ile Cys Arg Asn His Ile Met Asp Leu Cys Ile Glu Cys Gln Ala Asn Gln 60 |
| Db | 1498 | AACTGTGCCATCTGTAGGAATCACATTATGGATCTGCATAGAGTGCCAGGCTAACAG 1557 |
| Qy | 61 | Ala Ser Ala Thr Ser Glu Glu Cys Thr Val Ala Trp Gly Val Cys Asn His Ala Phe His 80 |
| Db | 1558 | GCCTCTGCCACATCAGAGGAGTGCACCGTAGCCTGGGAGTCTGCAATCATGCATTCCAT 1617 |
| Qy | 81 | Phe His Cys Ile Ser Arg Trp Leu Lys Thr Arg Gln Val Cys Pro Leu Asp Asn Arg Glu 100 |
| Db | 1618 | TTCCACTGTATCTCCGTTGGTTGAAGACCAGGCAGGTGTGCCCTAGACAACAGGGAG 1677 |
| Qy | 101 | Trp Glu Phe Gln Lys Tyr Gly His 108 |
| Db | 1678 | TGGGAGTTTCAGAAATATGGACAC 1701 |

RESULT 8
HUMYQ60A05
LOCUS HUMYQ60A05 497 bp mRNA linear PRI 29-AUG-1998
DEFINITION Homo sapiens full length insert cDNA clone YQ60A05.
ACCESSION AF085906
VERSION AF085906.1 GI:3483246
KEYWORDS FLI_CDNA.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 497)
AUTHORS Woessner,J., Tan,F., Marra,M., Kucaba,T., Yandell,M., Martin,J.,
Marth,G., Bowles,L., Wylie,T., Bowers,Y., Steptoe,M., Theising,B.,
Geisel,S., Allen,M., Underwood,K., Chappell,J., Person,B.,
Gibbons,M., Harvey,N., Pape,D., Chamberlain,A., Morales,R.,
Schurk,R., Ritter,E., Kohn,S., Swaller,T., Behymer,K., Hillier,L.,
Wilson,R. and Waterston,R.
TITLE Full Clone Sequencing of the Longest Available Member from Each
Unigene Cluster
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 497)
AUTHORS Waterston,R.
TITLE Direct Submission
JOURNAL Submitted (24-AUG-1998) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
COMMENT SUBMITTED BY:
Genome Sequencing Center
Department of Genetics
Washington University
St. Louis MO 63108, USA
<http://genome.wustl.edu/gsc>
<mailto:est@watson.wustl.edu>

NOTICE: This sequence represents the full insert of this cDNA. No attempt has been made to verify whether this corresponds to the full-length of the original mRNA from which it was derived. We have tried to obtain double-stranded, or double chemistry sequence across the entire clone, but potentially, there are areas in the sequence where this level of coverage was not achieved. Nevertheless, we are confident of the accuracy of this sequence as all regions of low quality, as defined by PHRAP (P. Green, in preparation), were visually inspected and edited accordingly. The consensus quality values for this sequence have been submitted separately.

SIMILARITY INFORMATION:
similar to *Caenorhabditis elegans* protein Z70757 (PID:g1262999)
ZK287.5

The location of this clone is unknown.

FEATURES Location/Qualifiers
source 1..497
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:200144"

```

        .           /clone_lib="Soares_fetal_liver_spleen_1NFLS"
misc_feature 3. .305
/note="similar to Caenorhabditis elegans protein Z70757
(PID:g1262999)"
misc_feature 18. .284
/note="similar to Caenorhabditis elegans protein U80449
(PID:g1707068)"
misc_feature 36. .302
/note="similar to Saccharomyces cerevisiae protein S66830
(PID:g2132017)"
misc_feature 42. .302
/note="similar to Schizosaccharomyces pombe protein Z98977
(PID:g2388937)"
misc_feature 51. .284
/note="similar to Caenorhabditis elegans protein Z46242
(PID:g559430)"
BASE COUNT      131 a      101 c      117 g      148 t
ORIGIN

```

Alignment Scores:

| | | | |
|------------------------|----------|---------------|-----|
| Pred. No.: | 2.87e-55 | Length: | 497 |
| Score: | 584.00 | Matches: | 101 |
| Percent Similarity: | 100.00% | Conservative: | 0 |
| Best Local Similarity: | 100.00% | Mismatches: | 0 |
| Query Match: | 94.81% | Indels: | 0 |
| DB: | 9 | Gaps: | 0 |

US-09-541-462B-2 (1-108) x HUMYQ60A05 (1-497)

| | | |
|----|---|--|
| Qy | 8 AspThrProSerGlyThrAsnSerGlyAlaGlyLysLysArgPheGluValLysLysTrp 27 | |
| Db | 3 GATACCCCGAGCGGCACCAACAGCGCGCGGGCAAGAAGCGCTTGAAAGTGAAAAAGTGG 62 | |
| Qy | 28 AsnAlaValAlaLeuTrpAlaTrpAspIleValValAspAsnCysAlaIleCysArgAsn 47 | |
| Db | 63 AATGCAGTAGCCCTCTGGGCTGGGATATTGTGGTTGATAACTGTGCCATCTGCAGGAAC 122 | |
| Qy | 48 HisIleMetAspLeuCysIleGluCysGlnAlaAsnGlnAlaSerAlaThrSerGluGlu 67 | |
| Db | 123 CACATTATGGATCTTGCATAGAACATGTCAAGCTAACCGAGCGTCCGCTACTTCAGAACAGAG 182 | |
| Qy | 68 CysThrValAlaTrpGlyValCysAsnHisAlaPheHisPheHisCysIleSerArgTrp 87 | |
| Db | 183 TGTACTGTCGCATGGGAGTCTGTAACCATGCTTTCACTTCCACTGCATCTCGCTGG 242 | |
| Qy | 88 LeuLysThrArgGlnValCysProLeuAspAsnArgGluTrpGluPheGlnLysTyrGly 107 | |
| Db | 243 CTCAAAACACGACAGGTGTCCATTGGACAAACAGAGAGTGGAAATTCCAAAAGTATGGG 302 | |
| Qy | 108 His 108 | |
| Db | 303 CAC 305 | |

RESULT 9

AX281690/c

LOCUS AX281690 5347 bp DNA linear PAT 02-NOV-2001
DEFINITION Sequence 99 from Patent WO0177389.

ACCESSION AX281690
 VERSION AX281690.1 GI:16608941
 KEYWORDS .
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1
 AUTHORS Shiffman,D., Somogyi,R., Lawn,R., Seilhamer,J.J., Porter,G.J.,
 Mikita,T. and Tai,J.
 TITLE Genes expressed in foam cell differentiation
 JOURNAL Patent: WO 0177389-A 99 18-OCT-2001;
 Incyte Genomics, Inc. (US)
 FEATURES Location/Qualifiers
 source 1. .5347
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /note="Incyte ID No: 364940.19"
 BASE COUNT 1489 a 1277 c 1094 g 1487 t
 ORIGIN

Alignment Scores:

| | | | |
|------------------------|----------|---------------|------|
| Pred. No.: | 5.74e-53 | Length: | 5347 |
| Score: | 573.00 | Matches: | 103 |
| Percent Similarity: | 99.04% | Conservative: | 0 |
| Best Local Similarity: | 99.04% | Mismatches: | 0 |
| Query Match: | 93.02% | Indels: | 1 |
| DB: | 6 | Gaps: | 0 |

US-09-541-462B-2 (1-108) x AX281690 (1-5347)

| | | |
|----|---|-----|
| Qy | 3 AlaAlaMetAspValAspThrProSerGlyThrAsnSerGlyAla-GlyLysLysArgPh | 22 |
| Db | 515 GCAGCGATGGATGTGGATACCCCGAGCGGCACCAACAGCGGCGCGGGCAAGAACGCCTT | 456 |
| Qy | 22 eGluValLysLysTrpAsnAlaValAlaLeuTrpAlaTrpAspIleValValAspAsnCy | 42 |
| Db | 455 TGAAGTAAAAAGTGGAATGCAGTAGCCCTCTGGCCTGGGATATTGTGGTTGATAACTG | 396 |
| Qy | 42 sAlaIleCysArgAsnHisIleMetAspLeuCysIleGluCysGlnAlaAsnGlnAlaSe | 62 |
| Db | 395 TGCCATCTGCAGGAACCACATTATGGATCTTGATAGAACATGCTAACCGAGCGTC | 336 |
| Qy | 62 rAlaThrSerGluGluCysThrValAlaTrpGlyValCysAsnHisAlaPheHisPheHi | 82 |
| Db | 335 CGCTACTTCAGAAGAGTGTACTGTCGCATGGGAGTCTGTAACCATGCTTTCACTTCCA | 276 |
| Qy | 82 sCysIleSerArgTrpLeuLysThrArgGlnValCysProLeuAspAsnArgGluTrpGl | 102 |
| Db | 275 CTGCATCTCTCGCTGGCTAAACACACGACAGGTGTCCATTGGACAAACAGAGAGTGGGA | 216 |
| Qy | 102 uPheGlnLys 105 | |
| Db | 215 ATTCCAAAAG 206 | |

RESULT 10

HTEST/c

LOCUS HTEST 3484 bp mRNA linear PRI 01-MAY-1995
 DEFINITION H.sapiens mRNA for testican.
 ACCESSION X73608
 VERSION X73608.1 GI:793844
 KEYWORDS testican.
 SOURCE Homo sapiens.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 3484)
 AUTHORS Alliel,P.M., Perin,J.P., Jolles,P. and Bonnet,F.J.
 TITLE Testican, a multidomain testicular proteoglycan resembling
 modulators of cell social behaviour
 JOURNAL Eur. J. Biochem. 214 (1), 347-350 (1993)
 MEDLINE 93285162
 PUBMED 8389704
 FEATURES Location/Qualifiers
 source 1..3484
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="lambda gt11"
 CDS 435..1754
 /codon_start=1
 /product="testican"
 /protein_id="CAA51999.1"
 /db_xref="GI:793845"
 /db_xref="SPTREMBL:Q08629"
 /translation="MPAIAVLAAAAAAWCFLQVESRHL DALAGGAGPNHGNFLDNDQWLSTVSQYDRDKYWNRFRDDYFRNWNPNKPFQALDPSKDPCLKVKCSHKVCTQDYQTALCVSRKHLLPRQKKGNVAQKHWVGPSNLVKCKPCPVAQSAMVCGSDGHSYTSKCKLEFHACSTGKSLATLCDGCPCLPEPEPPKHKKAERSACTDKELRNLASRLKDWF GALHEDANRVIKPTSSNTAQGRFDTSILPICKDSLGMFNKLDMNYDLLLDPSEINAIYLDKYEPCIKPLFNSCDSFKDGKLSNNEWCYCFQKPGGLPCQNE MNRIQKLSKGKSLLGAFIPRCNEEGYYKATQCHGSTGQCWCVDKYGNELAGSRKQGAVSCEEQETSGDFGS GGSV VLLDDLEYERELGPKDKEGKLRVHTRAVTEDDEDEDDEDDKEDEVGYIW"
 BASE COUNT 894 a 910 c 808 g 872 t
 ORIGIN

Alignment Scores:

| | | | |
|------------------------|----------|---------------|------|
| Pred. No.: | 4.45e-51 | Length: | 3484 |
| Score: | 554.00 | Matches: | 98 |
| Percent Similarity: | 100.00% | Conservative: | 0 |
| Best Local Similarity: | 100.00% | Mismatches: | 0 |
| Query Match: | 89.94% | Indels: | 0 |
| DB: | 9 | Gaps: | 0 |

US-09-541-462B-2 (1-108) x HTEST (1-3484)

| | |
|----|---|
| Qy | 3 Ala Ala Met Asp Val Asp Thr Pro Ser Gly Thr Asn Ser Gly Ala Gly Lys Lys Arg Phe 22 |
| Db | 296 GCAGCGATGGATGTGGATACCCCGAGCGGCACCAACAGCGGCGGGCAAGAACCGCTTT 237 |
| Qy | 23 Glu Val Lys Lys Trp Asn Ala Val Ala Leu Trp Ala Trp Asp Ile Val Val Asp Asn Cys 42 |
| Db | 236 GAAGTGAAAAAGTGGATGCAGTAGCCCTCTGGGCCTGGGATATTGTGGTTGATAACTGT 177 |
| Qy | 43 Ala Ile Cys Arg Asn His Ile Met Asp Leu Cys Ile Glu Cys Gln Ala Asn Gln Ala Ser 62 |

```

Db      176 |||||||GCCATCTGCAGGAACCACATTATGGATCTTGCATAGAATGTCAAGCTAACAGGCGTCC 117
Qy      63 AlaThrSerGluGluCysThrValAlaTrpGlyValCysAsnHisAlaPheHisPheHis 82
Db      116 |||||||GCTACTTCAGAAGAGTGTACTGTCGATGGGGAGTCTGTAACCATGCTTCACTTCCAC 57
Qy      83 CysIleSerArgTrpLeuLysThrArgGlnValCysProLeuAspAsnArgGlu 100
Db      56 |||||||TGCATCTCGCTGGCTAAAACACGACAGGTGTGCCATTGGACAAACAGAGAG 3

RESULT 11
AY099360
LOCUS      AY099360          306 bp      mRNA      linear      PRI 13-MAY-2002
DEFINITION Homo sapiens ZYP protein mRNA, partial cds.
ACCESSION  AY099360
VERSION    AY099360.1  GI:20502054
KEYWORDS   .
SOURCE     Homo sapiens.
ORGANISM   Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 306)
AUTHORS   Perin,J.-P., Seddiki,N., Charbonnier,F., Goudou,D., Belkadi,L.,
          Rieger,F. and Alliel,P.M.
TITLE      Genomic organization and expression of the ubiquitin-proteasome
          complex-associated protein Rbx1/ROC1/Hrt1
JOURNAL   Cell. Mol. Biol. (Noisy-le-grand) 45 (8), 1131-1137 (1999)
MEDLINE   20106778
PUBMED   10643962
REFERENCE  2 (bases 1 to 306)
AUTHORS   Alliel,P.M., Seddiki,N., Belkadi,L., Lecoeur,L. and Perin,J.P.
TITLE      Direct Submission
JOURNAL   Submitted (23-APR-2002) U488, INSERM, 80, rue du General Leclerc,
          Le Kremlin-Bicetre 94270, France
FEATURES   Location/Qualifiers
source     1. .306
          /organism="Homo sapiens"
          /db_xref="taxon:9606"
          /clone="zyp3"
          /sex="male"
          /tissue_type="whole brain"
          /dev_stage="26-week fetus"
          /note="caucasian"
          <1. .290
          /note="similar to RBX1/ROC1/HRT1"
          /codon_start=3
          /product="ZYP protein"
          /protein_id="AAM21718.1"
          /db_xref="GI:20502055"
          /translation="NSGASKKRFEVKKWNVALWAUDIVVDNCAICRNHIMDLCIECQ
          ANQASATSEECTVAWGVCNHAFFHCISRWLKTRQVCPLDNREWEFQKYGH"
CDS
BASE COUNT 84 a    70 c    80 g    72 t
ORIGIN


```

Alignment Scores:

Pred. No.: 3.29e-51 Length: 306

Score: 545.00 Matches: 94
 Percent Similarity: 98.95% Conservative: 0
 Best Local Similarity: 98.95% Mismatches: 1
 Query Match: 88.47% Indels: 0
 DB: 9 Gaps: 0

US-09-541-462B-2 (1-108) x AY099360 (1-306)

Qy 14 AsnSerGlyAlaGlyLysLysArgPheGluValLysLysTrpAsnAlaValAlaLeuTrp 33
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 3 AACAGCGCGCGAGCAAGAAGCGCTTGAAAGTGGAAATGCAGTAGCCCTCTGG 62

Qy 34 AlaTrpAspIleValValAspAsnCysAlaIleCysArgAsnHisIleMetAspLeuCys 53
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 63 GCCTGGGATATTGTGGTTGATAACTGTGCCATCTGCAGGAACCACATTATGGATCTTGC 122

Qy 54 IleGluCysGlnAlaAsnGlnAlaSerAlaThrSerGluGluCysThrValAlaTrpGly 73
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 123 ATAGAATGTCAAGCTAACCAAGCGTCCGCTACTTCAGAAGAGTGTACTGTCGCATGGGA 182

Qy 74 ValCysAsnHisAlaPheHisPheHisCysIleSerArgTrpLeuLysThrArgGlnVal 93
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 183 GTCTGTAACCATGCTTTCACTTCCACTGCATCTCGCTGGCTAAACACGACAGGTG 242

Qy 94 CysProLeuAspAsnArgGluTrpGluPheGlnLysTyrGlyHis 108
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 243 TGTCCATTGGACAACAGAGAGTGGGAATTCCAAAAGTATGGGCAC 287

RESULT 12

AY119265

LOCUS AY119265 1181 bp mRNA linear INV 16-JUN-2002
 DEFINITION Drosophila melanogaster SD23839 full insert cDNA.
 ACCESSION AY119265
 VERSION AY119265.1 GI:21430893
 KEYWORDS FLI_CDNA.
 SOURCE fruit fly.
 ORGANISM Drosophila melanogaster
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 Ephydriidea; Drosophilidae; Drosophila.
 REFERENCE 1 (bases 1 to 1181)
 AUTHORS Stapleton,M., Brokstein,P., Hong,L., Agbayani,A., Carlson,J.,
 Champe,M., Chavez,C., Dorsett,V., Dresnek,D., Farfan,D., Frise,E.,
 George,R., Gonzalez,M., Guarin,H., Kronmiller,B., Li,P., Liao,G.,
 Miranda,A., Mungall,C.J., Nunoo,J., Pacleb,J., Paragas,V., Park,S.,
 Patel,S., Phouanenavong,S., Wan,K., Yu,C., Lewis,S.E., Rubin,G.M.
 and Celniker,S.
 TITLE Direct Submission
 JOURNAL Submitted (07-JUN-2002) Berkeley Drosophila Genome Project,
 Lawrence Berkeley National Laboratory, One Cyclotron Road,
 Berkeley, CA 94720, USA
 COMMENT Sequence submitted by:
 Berkeley Drosophila Genome Project
 Lawrence Berkeley National Laboratory
 Berkeley, CA 94720
 This clone was sequenced as part of a high-throughput process to
 sequence clones from Drosophila Gene Collection 1 (Rubin et al.,

Science 2000). The sequence has been subjected to integrity checks for sequence accuracy, presence of a polyA tail and contiguity within 100 kb in the genome. Thus we believe the sequence to reflect accurately this particular cDNA clone. However, there are artifacts associated with the generation of cDNA clones that may have not been detected in our initial analyses such as internal priming, priming from contaminating genomic DNA, retained introns due to reverse transcription of unspliced precursor RNAs, and reverse transcriptase errors that result in single base changes. For further information about this sequence, including its location and relationship to other sequences, please visit our Web site (<http://fruitfly.berkeley.edu>) or send email to cdna@fruitfly.berkeley.edu.

| FEATURES | Location/Qualifiers |
|------------|--|
| source | 1. .1181 /organism="Drosophila melanogaster" /db_xref="taxon:7227" |
| gene | 1. .1181 /gene="Rocla" /db_xref="FLYBASE:FBgn0025638" |
| CDS | 162. .488 /gene="Rocla" /note="Longest ORF" /codon_start=1 /product="SD23839p" /protein_id="AAM51125.1" /db_xref="GI:21430894" /db_xref="FLYBASE:FBgn0025638" /translation="MEVDEDGYEVPSSSKGDKKRFEVKWNAVALWAUDIVVDNCAI CRNHIMDLCIECQANQASATSEECTVAWGVCNHAFHFCISRWLKTRQVCPLDNREWD FQKYGH" |
| BASE COUNT | 346 a 246 c 288 g 301 t |
| ORIGIN | |

Alignment Scores:

| | | | |
|------------------------|----------|---------------|------|
| Pred. No.: | 1.73e-49 | Length: | 1181 |
| Score: | 535.00 | Matches: | 95 |
| Percent Similarity: | 90.09% | Conservative: | 5 |
| Best Local Similarity: | 85.59% | Mismatches: | 5 |
| Query Match: | 86.85% | Indels: | 6 |
| DB: | 3 | Gaps: | 2 |

US-09-541-462B-2 (1-108) x AY119265 (1-1181)

| | | |
|----|--|-----|
| Qy | 3 AlaAlaMetAspValAsp-----ThrProSerGlyThrAsnSerGlyAla | 17 |
| | ::: ::: :::::: | |
| Db | 156 TCCACCATGGAAGTCGACGAGGATGGATACGAGGTTCCCTCCAGCAGCAGCAAGGGC--- | 212 |
| Qy | 18 GlyLysLysArgPheGluValLysLysTrpAsnAlaValAlaLeuTrpAlaTrpAspIle | 37 |
| | | |
| Db | 213 GATAAGAACCGCTTGAGGTGAAGAAGTGGAACGCCGTGGCTCTGTGGCCTGGGACATC | 272 |
| Qy | 38 ValValAspAsnCysAlaIleCysArgAsnHisIleMetAspLeuCysIleGluCysGln | 57 |
| | | |
| Db | 273 GTGGTGGACAACCTGCGCCATCTGCCGCAACCACATCATGGACTTGTGCATCGAGTGTCA | 332 |
| Qy | 58 AlaAsnGlnAlaSerAlaThrSerGluGluCysThrValAlaTrpGlyValCysAsnHis | 77 |

Db 333 GCGAACCAGGCCTCCGCCACTAGCGAGGAGTGCACCGTGGCCTGGGCCTCTGCAACCAC 392
 |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
 Qy 78 AlaPheHisPheHisCysIleSerArgTrpLeuLysThrArgGlnValCysProLeuAsp 97
 |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
 Db 393 GCCTTCCATTCCACTGCATCTCGCTGGCTAAAGACGCGCCAGGTATGCCCACTGGAC 452
 |||||||:|||||:|||||:|||||:
 Qy 98 AsnArgGluTrpGluPheGlnLysTyrGlyHis 108
 |||||||:|||||:
 Db 453 AACCGCGAGTGGGATTCCAGAAGTACGGCCAC 485

RESULT 13

AY086913

LOCUS AY086913 635 bp mRNA linear PLN 26-JUN-2002
 DEFINITION *Arabidopsis thaliana* clone 29408 mRNA, complete sequence.
 ACCESSION AY086913
 VERSION AY086913.1 GI:21405628
 KEYWORDS FLI_CDNA.
 SOURCE thale cress.
 ORGANISM *Arabidopsis thaliana*
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Rosidae; eurosids II; Brassicales; Brassicaceae; *Arabidopsis*.
 REFERENCE 1 (bases 1 to 635)
 AUTHORS Haas,B.J., Volfovsky,N., Town,C.D., Troukhan,M., Alexandrov,N.,
 Feldmann,K.A., Flavell,R.B., White,O. and Salzberg,S.L.
 TITLE Full-length messenger RNA sequences greatly improve genome
 annotation
 JOURNAL Genome Biol. (2002) In press
 REFERENCE 2 (bases 1 to 635)
 AUTHORS Brover,V., Troukhan,M., Alexandrov,N., Lu,Y.-P., Flavell,R. and
 Feldmann,K.
 TITLE Full-Length cDNA from *Arabidopsis thaliana*
 JOURNAL Unpublished
 REFERENCE 3 (bases 1 to 635)
 AUTHORS Brover,V., Troukhan,M., Alexandrov,N., Lu,Y.-P., Flavell,R. and
 Feldmann,K.
 TITLE Direct Submission
 JOURNAL Submitted (11-MAR-2002) Ceres, Inc, 3007 Malibu Canyon Road,
 Malibu, CA 90265, USA
 COMMENT This clone sequence is one of 5,000 Ceres full-length cDNAs made
 available to TIGR and Genbank. The following quality assessment of
 this set was done by comparison with known proteins: two percent of
 the clones are estimated to be 5'-truncated; less than one percent
 are 3'-truncated; approximately two percent represent alternative
 splice variants, including unspliced introns and spliced exons; one
 percent may contain premature stop codons; five percent may have
 frame shifts in a coding region. A sequence is considered to be
 5'-truncated if it lacks the translation initiation start (ATG). A
 sequence is considered to be 3'-truncated if it lacks the
 C-terminal end of the encoded protein. Please note that these cDNA
 sequences are derived from the Ws or LAer ecotypes and therefore
 may contain polymorphisms when compared to sequences from Col-0.
 Genset carried out the library production and sequencing of the
 full-length clones. Ceres, Inc. carried out the clustering of the
 5' sequences, selection of clones, and sequence assembly.
 FEATURES Location/Qualifiers

source 1. .635
/organism="Arabidopsis thaliana"
/db_xref="taxon:3702"
/clone="29408"
CDS 91. .447
/codon_start=1
/product="ring-box protein-like"
/protein_id="AAM64477.1"
/db_xref="GI:21592528"
/translation="MATLDSVTMIPAGEASSSVAASSSNKAKRFEIKKWSAVALWA
WDIVVDNCAICRNHIMDLCIECQANQASATSEECTVAWGVCNHAFFHCISRWLKTRQ
VCPLDNSEWEFQKYGH"
BASE COUNT 164 a 135 c 144 g 192 t
ORIGIN

Alignment Scores:

| | | | |
|------------------------|----------|---------------|-----|
| Pred. No.: | 1.24e-47 | Length: | 635 |
| Score: | 515.50 | Matches: | 92 |
| Percent Similarity: | 83.19% | Conservative: | 7 |
| Best Local Similarity: | 77.31% | Mismatches: | 9 |
| Query Match: | 83.69% | Indels: | 11 |
| DB: | 8 | Gaps: | 2 |

US-09-541-462B-2 (1-108) x AY086913 (1-635)

| | |
|---|---|
| Qy | 1 MetAlaAlaAlaMetAspValAspThr-----ProSerGly----- 12 |
| ::: :::: ::: | |
| Db | 88 TTAATGGCGACTCTAGACTCCGACGTTACCATGATTCTGCCGGAGAACCTCCAGCAGC 147 |
| Qy | 13 -----ThrAsnSerGlyAlaGlyLysLysArgPheGluValLysLysTrpAsnAla 29 |
| ::::: ::: ::: | |
| Db | 148 GTAGCCCGCTCGTCTTCCAACAAGAAAGCTAACGATTGAAATTAGAAGTGGAGGCC 207 |
| Qy | 30 ValAlaLeuTrpAlaTrpAspIleValValAspAsnCysAlaIleCysArgAsnHisIle 49 |
| | |
| Db | 208 GTTGCTCTCTGGGCTTGGGATATCGTTGACAACGTGCGATCTGCAGAAACCACATC 267 |
| Qy | 50 MetAspLeuCysIleGluCysGlnAlaAsnGlnAlaSerAlaThrSerGluGluCysThr 69 |
| | |
| Db | 268 ATGGATCTTGTATCGAGTGTCAAGCTAACGCTTCCACTTCACTGCATCAGCAGATGGCTAAAG 327 |
| Qy | 70 ValAlaTrpGlyValCysAsnHisAlaPheHisPheHisCysIleSerArgTrpLeuLys 89 |
| | |
| Db | 328 GTAGCTGGGGGGTTGCAATCACGCCTTCCACTTCACTGCATCAGCAGATGGCTAAAG 387 |
| Qy | 90 ThrArgGlnValCysProLeuAspAsnArgGluTrpGluPheGlnLysTyrGlyHis 108 |
| | |
| Db | 388 ACTCGTCAAGTTGTCCATTGGATAACAGTGAGTGGAGTTCAGAAATATGGTCAC 444 |

RESULT 14

AY072430

| | | | | | |
|-------------------|--|--------|------|--------|-----------------|
| LOCUS | AY072430 | 666 bp | mRNA | linear | PLN 21-JAN-2002 |
| DEFINITION | Arabidopsis thaliana ring-box protein - like (At5g20570) mRNA, complete cds. | | | | |
| ACCESSION | AY072430 | | | | |
| VERSION | AY072430.1 GI:18252990 | | | | |
| KEYWORDS | FLI_CDNA. | | | | |

SOURCE *Arabidopsis thaliana*.
 ORGANISM *Arabidopsis thaliana*
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicots; core eudicots;
 Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 REFERENCE 1 (bases 1 to 666)
 AUTHORS Nguyen,M., Karlin-Neumann,G., Southwick,A., Lam,B., Miranda,M.,
 Palm,C.J., Bowser,L., Jones,T., Banh,J., Carninci,P., Chen,H.,
 Cheuk,R., Chung,M.K., Hayashizaki,Y., Ishida,J., Kamiya,A.,
 Kawai,J., Kim,C., Lin,J., Liu,S.X., Narusaka,M., Pham,P.K.,
 Sakano,H., Sakurai,T., Satou,M., Seki,M., Shinn,P., Yamada,K.,
 Shinozaki,K., Ecker,J., Theologis,A. and Davis,R.W.
 TITLE Direct Submission
 JOURNAL Submitted (02-JAN-2002) DNA Sequencing and Technology Center,
 Stanford University, 855 California Avenue, Palo Alto, CA 94304,
 USA
 COMMENT e-mail for correspondence: arab@sequence.stanford.edu

RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN *Arabidopsis* Full-Length cDNA'): Seki,M., Narusaka,M., Ishida,J., Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J., Hayashizaki,Y. and Shinozaki,K.

The Salk, Stanford, PGEC (SSP) Consortium members carried out the sequencing and annotation of the RAFL cDNAs: Nguyen,M., Southwick,A., Karlin-Neumann,G., Lam,B., Miranda,M., Palm,C.J., Bowser,L., Jones,T., Banh,J., Chen,H., Cheuk,R., Chung,M.K., Kim,C., Lin,J., Liu,S.X., Pham,P.K., Sakano,H., Shinn,P., Yamada,K., Ecker,J., Theologis,A. and Davis,R.W.

Nguyen,M. (SSP/Stanford) and Seki,M. (RIKEN GSC) contributed equally to this work. Shinozaki,K. (RIKEN GSC) and Davis,R.W. (SSP/Stanford) contributed equally to this work as PIs.

FEATURES Location/Qualifiers
 source 1. .666
 /organism="Arabidopsis thaliana"
 /db_xref="taxon:3702"
 /chromosome="5"
 /clone="RAFL09-78-L12"
 /note="This clone is in pBluescript
 ecotype: Columbia"
 gene 1. .666
 /gene="At5g20570"
 CDS 88. .444
 /gene="At5g20570"
 /codon_start=1
 /product="ring-box protein - like"
 /protein_id="AAL62422.1"
 /db_xref="GI:18252991"
 /translation="MATLDSVTMIPAGEASSSVAASSSNKKAKRFEIKKWSAVALWA
 WDIVVDNCAICRNHIMDLCIECQANQASATSEECTVAWGVCNHAFFHCISRWLKTRQ
 VCPLDNSEWEFQKYGH"

BASE COUNT 181 a 137 c 149 g 199 t
 ORIGIN

Alignment Scores:

Pred. No.: 1.3e-47 Length: 666
 Score: 515.50 Matches: 92
 Percent Similarity: 83.19% Conservative: 7
 Best Local Similarity: 77.31% Mismatches: 9
 Query Match: 83.69% Indels: 11
 DB: 8 Gaps: 2

US-09-541-462B-2 (1-108) x AY072430 (1-666)

Qy 1 MetAlaAlaAlaMetAspValAspThr-----ProSerGly----- 12
 ::|:|||:::||| |||:|||:
 Db 85 TTAATGGCGACTCTAGACTCCGACGTTACCATGATTCCCTGCCGGAGAACCTCCAGCAGC 144

 Qy 13 -----ThrAsnSerGlyAlaGlyLysLysArgPheGluValLysLysTrpAsnAla 29
 :::::::||| |||||||:|||:|||:|||:|||:
 Db 145 GTAGCCCGCTCGTCTTCCAACAAAGAAAGCTAACGCGATTGAAATTAAGAAGTGGAGCGCC 204

 Qy 30 ValAlaLeuTrpAlaTrpAspIleValValAspAsnCysAlaIleCysArgAsnHisIle 49
 |||||||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:
 Db 205 GTTGCTCTCTGGGCTTGGGATATCGTTGTTGACAACGTGCGATCTGCAGAAACCACATC 264

 Qy 50 MetAspLeuCysIleGluCysGlnAlaAsnGlnAlaSerAlaThrSerGluGluCysThr 69
 |||||||:|||:|||:|||:|||:|||:|||:|||:|||:|||:
 Db 265 ATGGATCTTGTATCGAGTGTCACTCAGGCCAGTGCACAAAGTGAAGAGTGCAC 324

 Qy 70 ValAlaTrpGlyValCysAsnHisAlaPheHisPheHisCysIleSerArgTrpLeuLys 89
 |||||||:|||:|||:|||:|||:|||:|||:|||:
 Db 325 GTAGCTGGGGGTTGCAATCACGCCTTCCACTTCAGCAGATGGCTAAAG 384

 Qy 90 ThrArgGlnValCysProLeuAspAsnArgGluTrpGluPheGlnLysTyrGlyHis 108
 |||||||:|||:|||:|||:|||:
 Db 385 ACTCGTCAAGTTGTCCATTGGATAACAGTGAGTGGAGTTCAGAAATATGGTCAC 441

RESULT 15

AY052401

LOCUS AY052401 357 bp mRNA linear PLN 16-OCT-2001
 DEFINITION Arabidopsis thaliana ring box-1-like protein mRNA, complete cds.
 ACCESSION AY052401
 VERSION AY052401.1 GI:16186264
 KEYWORDS
 SOURCE Arabidopsis thaliana.
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 REFERENCE 1 (bases 1 to 357)
 AUTHORS Okresz,L.
 TITLE Direct Submission
 JOURNAL Submitted (25-AUG-2001) Institute of Plant Biology, Biological
 Research Center, Temesvari krt. 62, Szeged H-6701, Hungary
 FEATURES Location/Qualifiers
 source 1. .357
 /organism="Arabidopsis thaliana"
 /db_xref="taxon:3702"
 CDS 1. .357
 /note="RBX1-2"
 /codon_start=1

/product="ring box-1-like protein"
/protein_id="AAL13435.1"
/db_xref="GI:16186265"
/translation="MATLDSDVTMIPAGEASSSVAASSSNKKAKRFEIKKWSAVALWA
WDIVVDNCAICRNHIMDLCIECQANQASATSEECTVAWGVCNHAFFHCISRWLKTRQ
VCPLDNSEWEFQKYGH"

BASE COUNT 91 a 83 c 94 g 89 t
ORIGIN

Alignment Scores:

| | | | |
|------------------------|----------|---------------|-----|
| Pred. No.: | 8.67e-48 | Length: | 357 |
| Score: | 514.50 | Matches: | 92 |
| Percent Similarity: | 83.76% | Conservative: | 6 |
| Best Local Similarity: | 78.63% | Mismatches: | 8 |
| Query Match: | 83.52% | Indels: | 11 |
| DB: | 8 | Gaps: | 2 |

US-09-541-462B-2 (1-108) x AY052401 (1-357)

| | |
|---|--|
| Qy | 3 AlaAlaMetAspValAspThr-----ProSerGly----- 12 |
| :::: : | |
| Db | 4 GCGACTCTAGACTCCGACGTTACCATGATTCCCTGCCGGAGAACCTCCAGCAGCGTAGCC 63 |
| Qy | 13 ---ThrAsnSerGlyAlaGlyLysLysArgPheGluValLysLysTrpAsnAlaValAla 31 |
| ::::: : : : : : : : : : | |
| Db | 64 GCGTCGCTTCCAACAAGAAAGCTAACGATTGAAATTAAAGAAGTGGAGCGCCGTTGCT 123 |
| Qy | 32 LeuTrpAlaTrpAspIleValValAspAsnCysAlaIleCysArgAsnHisIleMetAsp 51 |
| : : : : : : : : : : : : : | |
| Db | 124 CTCTGGGCTTGGGATATCGTTGACAACGTGCGATCTGCAGAAACCACATCATGGAT 183 |
| Qy | 52 LeuCysIleGluCysGlnAlaAsnGlnAlaSerAlaThrSerGluGluCysThrValAla 71 |
| : : : : : : : : : : : : | |
| Db | 184 CTTTGTATCGAGTGTCAAGCTAACGCCAGTGCCACAAGTGAAGAGTGCACGTGACTGTAGCT 243 |
| Qy | 72 TrpGlyValCysAsnHisAlaPheHisPheHisCysIleSerArgTrpLeuLysThrArg 91 |
| : : : : : : : : : : : | |
| Db | 244 TGGGGGGTTTGCAATCACGCCCTCCACTTCACTGCATCAGCAGATGGCTAAAGACTCGT 303 |
| Qy | 92 GlnValCysProLeuAspAsnArgGluTrpGluPheGlnLysTyrGlyHis 108 |
| : : : : : : : : : : : | |
| Db | 304 CAAGTTGTCCATTGGATAACAGTGAGTGGAGTTCAGAAATATGGTCAC 354 |

Search completed: April 4, 2003, 17:52:09
Job time : 1471 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: April 4, 2003, 16:34:12 ; Search time 219 Seconds
(without alignments)
1110.575 Million cell updates/sec